DOCUMENT RESUME

ED 179 554 TM 009 749

AUTHOR Moyer, Kerry L.
TITLE Evaluative Research Systems for Educational Policy.
PUB DATE 6 Nov 78
NOTE 16p.; Paper presented at the annual meeting of the Concerned Leaders in Educational Administration and Research (2nd, Alexandria, VA, November 6, 1978)

EDRS PRICE MF01/PC01 Plus Postage.
DESCRIPTORS Affective Objectives; Change Strategies; *Conceptual Schemes; Data Collection; Decision Making; *Educational Assessment; *Educational Policy; Emotional Response; Evaluation; *Evaluation Criteria; *Evaluation Methods; Evaluators; Policy Formation; Problem Solving; Research Criteria; *Research Methodology; Research Problems; Thought Processes

ABSTRACT

Four evaluation methods for determining educational policy are objectivism; subjectivism; emotive-imperative; and instrumentalism. For the objectivist, the rationale for undertaking an evaluation is based on empirical verification or non-verification of claims of existing policies or activities. Secondly, the objectivist is interested in questioning what should exist in education. Objectivists may be disinterested in the agency of policy making decisions, and are mainly concerned with precision, methodological accuracy, and the reliability of the evidence. The subjectivist is interested in the affective domain of feelings, attitudes, and interests, and many use the questionnaire as a methodological tool. The subjectivist is concerned with the value of educational products in terms of approval or interest. For emotive-imperative theorists, in contrast to the subjectivists, praise or condemnation of programs is based on individual feelings, and emotive evaluation is used to change attitudes of policy makers. Although data collection is involved, results are presented as expressions of feelings. Instrumentalists are involved in influencing decisions which resolve problems practically. Instrumental evaluations are related to specific programs. These evaluations are associated with operations performed during the implementation of programs or policies. Educational policy is considered secondary to educational practice. (MH)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

U.S. DEPARTMENT OF HEALTH, EDUCATION & WELFARE
NATIONAL INSTITUTE OF EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

ED179554

Evaluative Research Systems for Educational Policy

Kerry L. Moyer

Division of Research
Pennsylvania Department of Education

November 1978

A paper presented at the Second Annual Conference of Concerned Leaders in Educational Administration and Research (CLEAR), formerly The Research and Development Association for Education (TRADAE), in Alexandria, Virginia, November 6, 1978

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

Kerry L. Moyer

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

TM009 749

Kerry Moyer
Research Associate
Division of Research
Pennsylvania
Department of
Education

EVALUATIVE RESEARCH SYSTEMS FOR EDUCATIONAL POLICY

It is possible to distinguish two types of activities under the topic of educational evaluation: (1) the collection of data and the making of judgments concerning educational activities, and (2) the discussion, study, or implementation of research evaluation findings as educational policy. Evaluating educational processes and products has always accompanied systems of schooling--be it the making of immediate, necessary judgments on a day-to-day basis in the classroom or "formal" experimental studies involving substantial resources and time requirements. Both activities represent the position that, at least in public schools, it is better to base educational decisions on evidence than prejudice. If day-to-day classroom data observations could sufficiently meet the information needs of the public, government, and educational institutions, there would be little argument or need for the more complex experimental or quasi-experimental methodological techniques commonly associated with the term "educational evaluation." In effect, collection of evidence has necessitated evaluation schemes and methodologies as complex as the educational or social systems to which they are applied. The extent to which these methodologies are available and adequate determines the usefulness of research findings to the policy maker.

The universe of evaluation activities has three main divisions: definition, rationale, and methods. The first set of evaluation questions addresses the topic of definition: What do the basic evaluation terms ("successful," "significant difference," "better than," "marginal") mean in everyday educational practice? What effects do these terms have on educational decision makers for whom they are intended? What other meanings might these words convey? Can the terms and conditions of the educational evaluation be understood by all interested audiences? Can certain vaguenesses or potential misunderstandings be explained satisfactorily in time to avoid such happenings? Do these terms represent actual observations or assertions and beliefs?

The principal questions in evaluation concerning rationale are the following: Is the study prescriptive or descriptive in nature? What is the role of the evaluator and the role of the policy maker in evaluation activities? What decisions or policy will be based on collected evidence? What was the actual motivation for the initiation of the evaluation study?

Evaluation questions concerning method include: Can the evaluation study be said to be internally and/or externally valid? What rules of inference were used? Are judgments deduced from sets of empirical observations? What is the validity and generalizability of measurement techniques employed? Does the study use logical steps to justify statistical tests? How well were mathematical assumptions and philosophies addressed? How efficient was this type of method? Are research findings useful to policy makers?

A further discussion of only one area, rationale, of educational evaluation follows. Carefully examining the rationale of an educational evaluation serves three important functions:

(1) As an aid in interpretation. It is possible for more than one set of values to be functioning in an educational evaluation. Most educational activities and projects have specific pre-stated values (be they goals, objectives, or intended outcomes) which may be flexible or inflexible as the program progresses. The values inherent in evaluation activities may or may not agree exactly with those of the policy makers. More commonly, a shift of values occurs as continuing educational services progress. Such shifts impose serious dilemmas on the evaluator; when such changes are radical (or fundamental) original studies become useless--yet the evaluator cannot dictate or change educational practice which seems proper simply to preserve existing methodology. Original study plans which have been retained despite later shifts in educational goals are subject to problems of interpretation. The policy maker, after identifying such conditions, is better prepared to judge results of evaluation activities as reported and to use observations collected as a basis for sound policy.

(2) As a measure of practical criteria. The term "practical criteria" is used here as described by Stufflebeam (1971) and includes: scope, relevance, importance, credibility, timeliness, and pervasiveness. One may use logical arguments or logical validation for definitions and rival hypotheses to judge methodological adequacies of an evaluation study. The rationale may, in like manner, be examined by careful study of the practicality of an educational evaluation. One very effective test of the usefulness of an evaluation is the amount of congruence between the information provided by the evaluator and the information needs of the policy maker. The better the congruence of research findings with policy needs the more useful the study.

(3) As a basis for increasing theory-oriented educational studies. Many evaluation studies are created from a methodological perspective; study parameters are often dictated by scientific considerations. Frequently, the idea producers of educational programs and services are the policy makers themselves; the scientific element is added by cooperating methodologists. Such evaluation attempts begin as theory-oriented studies and are finally produced as scientific studies developed through ad hoc procedures. Attempts at "data snooping" after the fact generally stray from educational interests. The rationale for any educational evaluation will indicate the educational theory, if any, underlying the educational service. Classifying evaluation studies by rationale may prove to be a method of combining the results of these many studies into more coherent and substantive bases for education policy decisions. An indirect result of increased educational theory study would be an expanded and established foundation of data which will greatly aid the educational decision maker.

The extent of this lack of theory-based evaluation studies is documented by Campbell and Stanley (1968, page 21):

> At present, there seem to be two main types of "experimentation" going on within schools: (1) research "imposed" upon the school by an outsider, who has his own ax to grind and whose goal is not immediate action (change) by the school, and (2) the so-called "action" researcher, who tries to get teachers themselves to be "experimenters," using that word quite loosely. The first researcher gets results that may be rigorous but not applicable. The latter gets results that may be highly applicable but probably not "true" because of extreme lack of rigor in the research.

What evaluators are seeking in the above example are measures of program outcomes. According to Fitz-Gibbon and Morris (1975, page 1):

> A theory-based evaluation of a program is one in which the selection of program features to evaluate is determined by an explicit conceptualization of the program in terms of a theory, a theory which attempts to explain how the program produces the desired effects.

When the theory or philosophy serving as the base for an educational program is identified early in the evaluative process the choice of variables of interest is based more on the needs of the policy makers and less on opinion and tradition. Variables of general importance in educational policy have by no means been identified; the composite results of many single educational evaluations may determine those variables which more need to be considered in the making of educational policy.

The rationale of an educational evaluation is premised upon the policy needs of the many participants of the evaluation activity. The greater the documentation of policy needs involved in the evaluation process, the more universally useful the observations gathered. The evaluation systems discussed generally to this point will now be identified, defined, and discussed in more detail. The following portion of this paper is devoted to four specific evaluation systems: objectivism, subjectivism, emotive-imperative, and instrumentalism. Each system provides the basis for policy decisions based upon differing philosophies. These terms are borrowed from meta-ethical and philosophical sources; however, their applicability to evaluation is very high and provides a framework for the critique of present evaluation studies and an outline for future studies. Each of these systems can be studied as an entity in itself; all four systems boast a considerable literature.

Objective Approaches to Educational Policy

Much of the present acceptance of the theory of objectivism can be traced to the writing of psychometricians and the outcomes of their work early in this century. Basically, this theory follows the now famous thought of Edward L. Thorndike (1918, page 16): "Whatever exists at all, exists in some amount. To know it thoroughly involves knowing its quantity as well as its quality." Scientific truth is found by determining the measurable occurrence or absence of a property of interest in the experimental units.

In general, the basis of educational policy might be translated into these questions: "What things ought to exist in education?" and "What actions best bring about what should exist?" It is apparent that if these questions could not be answered analytically there would be no need for objective educational evaluation. However regardless of how difficult it is to answer the first question through empiricism, the second question should not be addressed through other means in many cases.

Frequently the most difficult part of an evaluation from the objectivist point of view is the separation of these two questions. George E. Moore (1903, page vii), while discussing meta-ethics, has discussed this problem in a manner beneficial for evaluators:

> It appears to me that in Ethics, as in all other philosophical studies, the difficulties and disagreements, of which its history are full, are mainly due to a very simple cause: namely to the attempt to answer questions, without first discovering precisely what question it is which you desire to answer... At all events, philosophers seem, in general, not to make the attempt; and, whether the consequence of this omission or not, they are constantly endeavoring to prove that "Yes" or "No" will answer questions, to which neither answer is correct, owing to the fact that what they have been before their minds is not one question, but several, to some of which the true answer is "No," to others "Yes."

It will be beneficial to call the first question ("What things ought to exist in education?") the ends of education and the second question ("What actions best bring about what should exist?") the means of education. To be sure, things which are empirically determined to be the best "means" of education also have other properties such as expensive, time-consuming, or inefficient which must also be considered by the policy maker.

Many programs stem from the question of "What ought to exist in education," not the least of which is compensatory education. Compensatory education serves as a particularly useful example owing to the substantial involvement of evaluators with these programs. Studying the means of education has been a favorite approach to compensatory education by evaluators. This approach is particularly appealing to evaluators whose training is basically in research methodology; the compare/contrast approach with hypotheses supported by statistical techniques is direct and comfortable. This supports well the picture of the third party evaluator, objective, empirical, and disinterested, whose research ability is quite independent of particular educational or social contexts. Too frequently the result of this relationship is the attempt to base educational policy upon the advantage or pleasure of the educational evaluator. "This theory comes to grief over the fact, which stares us in the face, that a great part of duty consists in an observance of the rights and a furtherance of the interest of others, whatever the cost to ourselves may be" (Ross, 1930, page 16).

The objectivist as educational evaluator may not be directly concerned with the policy decisions of an educational institution or government agency. The evaluator's purpose is to supply numerical evidence of the desired outcomes of a specific educational attempt. Most evaluations are motivated and supported by government agencies; in these instances, evaluators deal not only with educational policies, but also with government policies and the broad social forces which brought such social or educational programs into being. Public schools and programs serve the public; the relationship of program participant and program implementation may not be the dominant policy concern. Ross (1930, page 21) identifies such prima facie duties of agents and individuals:

1. Duties based on previous acts of individuals

    a. Promises for services (fidelity)
    b. Correcting problems of wrong-doing (reparation)

2. Duties based on services from others (gratitude)
3. The distribution of happiness, wealth, rights, etc. (justice)
4. Improving the conditions of others (beneficence)
5. Services allowing improvement of conditions (self-improvement)
6. Other duties collectively which do not injure others

When an educational or social agency implements a program based on one or more of the above duties it falls upon the objectivist evaluator to determine empirically the amount of outcome of resulting acts. This is to be done without consideration of public or participant approval, opinion or support. The evaluation results are independent of and without reference to any human being; evaluation is "a matter of truth and falsity, and to say that a judgment is true obviously means something different from the statement that it is thought to be true" (Westermarck, 1960, page 3).

From the early psychometrists to contemporary research methodologists the theory of objective evaluation has enjoyed considerable influence--an influence which transferred to educational evaluation when methodologists directed attention toward that activity. Strangely, and in contrast to most other social sciences as sociology and anthropology, educational evaluation began with large scale designs and, often, national scope which seemed to require complex statistical techniques.

The rationale of an evaluation study, then, for the objectivist evaluator is the empirical verification of the truth or falsity of the claims of existing educational activities or policy. Secondarily, questions of what should exist in education and the amount of meeting of prima facie duties of the agent are addressed. Objectivist evaluators may be disinterested parties to the policy agency to preserve objectivity, reliability, and truthfulness of evidence collected. While not all objectivist evaluators subscribe to the disinterested party procedure, the general concern of research validity is precision and methodological accuracy and not questions of policy directly.

Subjectivism

The subjective educational evaluator studies people's feelings, attitudes and interests. Here the questions ("What ought to exist in education?" and "What actions best bring about what should exist?") are answered as approval or disapproval. The subjectivist argues that the means and ends of education are not measured in terms of purely objective properties; the "success" of an educational program is measured by the tendency of the program to arouse in people certain emotions. "Furthermore, unless a person felt such an emotion (or had such a tendency to feel it), the object would be neither good nor bad. Thus, it is people's feelings that make things good or bad" (Taylor, 1963, page 58).

The subjective evaluator freely admits the exchanges between the evaluator and the policy maker; indebtedness, gratitude, loyalty and even defensiveness may become an integral part of an educational program. Programs which give children great joy, for instance, make more difficult negative judgments in other areas of concern. "Whatever emotions may follow moral judgments, such judgments could never have been pronounced unless there had been moral emotions in somebody antecedent to them" (Westermarck, 1960, page 90). Of course, empiricism and intellectual inquiry may play substantial roles in the development of studies of approval or disapproval.

The relative strength of the objective and subjective influences on educational evaluation seems to be cyclic; both exist side-by-side, but their use in policy formation chang s with social conditions. With recent innovative educational programs "belief in an objective reason gave way to cultural relativity" (Blanshard, 1962, page 40). While objectivists were interested in empirically derived differences among observations, the subjectivist focused on the underlying identities of those differences. Further acceptance of subjective theory by some evaluators accompanied the realization that empirical evidence provided the policy maker was only one source of input to potential decisions. Lepley (1944, page 95) states well this condition:

> We go, perhaps, to the heart of the differences between what are commonly regarded as typical instances of knowing and valuation in their wider as well as in their more restricted meanings when we recognize that in fact finding--especially in scientific procedures--all desires are (ideally) excluded except the master desire for the truth, ...whereas in valuation, if suggestion is to be fruitful and decision adequate and far-seeing, all relevant desires must be brought into the arena of deliberation and given free play with relation to each other... Moreover, the decision about fact hinges entirely or predominantly upon perception (sensory or mental) of what is; but the decision regarding the valuable or the right course of action is ultimately determined wholly or more largely by the wish, preference, or purpose which wins in competition with other desires.

A favorite methodological tool of the subjectivist is the questionnaire. However, the subjectivist may develop an evaluation study based completely on subjective approaches; that is, the definition, rationale and methodology of the study may be all intuitive. The area of subjective probabilities (for instance, Kyburg & Smokler, 1964; Byrne, 1968) is one approach to subjective theory and its application to evaluation. Both subjectivist and objectivist study the same problems and the same educational programs; it is the emphasis that differs.

> We must find some connection between this abstract entity which satisfies certain mathematical stipulations and the pragmatic content, the real meaning, of the important statements of scientific and social intercourse which contain the word "probability" or one of its synonyms (Byrne, 1968, page 3).

A major division of subjective theory important to evaluators is the school which defines policy in terms of interest. To define what is valuable in an educational program the evaluator defines what is of interest.

> According to the definition of value here proposed, a thing--any thing--has value, or is valuable, in the original and generic sense when it is the object of an interest--any interest (Perry, 1968, page 2).

Things of interest in an educational setting are frequently the goals or objectives of a specific program. "Or, a thing is an object of interest when its being expected induces actions looking to its realization or nonrealization" (Perry, 1968, page 2). Here the evaluator is interested in verifying the value of program objectives as well as the ability of a program (or individual) to achieve intended outcomes. Because interests exist in individuals (and because interests are brought to the school as well as developed there) the value of educational programs must be measured in subjective terms. This is the task of the subjectivist evaluator.

Our previous example of compensatory education remains appropriate here; compensatory education has critics pro and con, many emotionally involved with issues. Empirical research which counters political and popular approval (such as that enjoyed by Head Start and Follow Through) is more likely to be attacked on emotional levels than scientific levels. This has been the case. Such research seems prone to attack on two fronts: (1) the variables studied were not of as great importance to the public as those variables not included (or superficially included) responsible for the program's popularity, and (2) the emotional involvement of the agents and the evaluators with the competency of reported findings. As an example of the first type of criticism, reference the following description of Follow Through evaluations:

> When the studies were designed, criterion measures representing common objectives of the various models having reasonably adequate reliability and validity were selected for inclusion in the evaluations. This lead to a heavy weighting of instruments assessing cognitive development... In general, the Planned Variation evaluations focus attention primarily on the limited range of program outcomes related to cognitive achievement (Bissell, 1973, page 70).

As an example of the second type of criticism, reference a statement from the first paragraph of The National Impact Study of Head Start:

> I was a consultant involved with the effort and, while I believe there are flaws in it, I also believe that the flaws are not critical and that the study's general conclusions are sound. None of the arguments I have yet seen has changed my mind about this (White, 1973, page 163).

Evaluation research differs from classical research in that questions of educational policy are always involved. The extent that these values influence both the evaluator's actions and the decisions of educational clients are of direct interest to the subjective educational evaluator. The value of an educational program is found in its scope of approval or its congruence with interests (be they governmental, social or personal). Because this approach deals more effectively with political realities and there is a growing dissatisfaction with the ability of objective research to deal with social reality, the subjective evaluator is approached by educational agents. "Program people are also alienated by the research armamentarium that forces them either to lie or to admit they are guilty of unpardonable sins such as lack of randomization, inadequate control groups, and even worse, having permitted the program to change in ways that were not made explicit at the beginning" (Guttentag, 1973). The subjectivist as educational evaluator, then, is concerned with the value of educational products and processes in terms of approval or interests.

Emotive-Imperative

The emotive-imperative theorists argue that objectivism and subjectivism are of little use because it is impossible to verify empirically what is true or false; such statements are actually functions of emotion. The emotivists believe that statements are justified, not proven, through emotional displays. The imperativists argue that behavior is based upon emotional responses triggered by statements internal or external. Generally, both schools of thought are combined to create the emotive-imperative theory. Evaluative results, despite the empiricism involved in data collection, are more expressions of feelings--and there is no way to verify the validity of feelings.

> All that one may legitimately enquire in this connection is, "What are the moral habits of a given person or group of people, and what causes them to have precisely those habits and feelings?" And this inquiry falls wholly within the scope of the existing social sciences (Ayer, 1952, page 31).

Emotive-imperative theory differs from the subjective theory in two important ways: (1) Praise or condemnation of educational programs is based solely on individual feelings and (2) the purpose of emotive evaluation is to directly change the feelings of the educational policy makers. Stevenson (1944, page 26) presents the method of the emotive-imperative evaluator:

> The model for "This is good" consists of the conjunction of (a) "I approve of this," and (b) "Do so as well." If a proof is positive for (a) and (b) taken separately, then and only then will it be possible for their conjunction. So let us see what can be done with sentences separately.
>
> Sentence (a) offers no trouble. It makes an assertion about the speaker's state of mind, and like any psychological statement, is open to empirical confirmation or discomfirmation, whether introspective or behavioristic.
>
> Sentence (b), however, raises a question. Since it is an imperative, it is not open to proof at all. What is it like to prove a command? If we told a person to close the door, and received the reply, "Prove it" should we not, to speak mildly, grow somewhat impatient?

Although imperatives cannot be proved, evidence is presented in an emotive-imperative evaluation which supports imperatives. An imperative is frequently met by the question "Why?" If the evaluation can satisfy reasons of "Why?" to the decision maker, the goal of changing the actions or attitudes of the policy maker has been met. For instance, the opening statement of an article by Jensen (1969) is: "Compensatory education has been tried and it apparently has failed." The remaining 123 pages are devoted to the reasons the reader should agree with Jensen. For a section of readers the question of "Why" will be answered; for others the question is not answered satisfactorily. But the fact remains that, even after the introduction of voluminous and complex studies, the disagreement and emotionalism continues, perhaps at an accelerating rate. This is the advantage of emotive-imperative evaluations: the simple introduction of empirical evidence leaves much of the emotionalism untouched.

There are several reasons for the existence of emotive-imperative evaluations: (1) research to support personal beliefs, (2) evaluations which command or influence the very preservation of educational programs, and (3) the use of scientific authority and appeal to support personal positions of educational theory. An advantage of this approach to other forms of evaluation research is the impossibility of making mistakes; evaluation is the expression of a favorable feeling not subject to truth or falsehood. Its usefulness is the degree to which the policy maker has been changed in some planned manner.

> Value involves, I think, at least the potentiality of choice, and it therefore attaches much more plausibly to future states of the world than to states of mind. A value is a future fact, selected from among a set of alternative, and mutually exclusive, future facts, and marked with an imperative. The chosen fact may not come about; in such a case we speak of the value as unrealized (Caws, 1967, page 61).

Obviously, the emotive-imperative evaluation is best suited to emotional issues or objective studies where factual remarks will arouse strong feelings and attitudes. A strong emotive-imperative evaluation study may influence policy makers in a comparatively uniform way, thus reducing conflict or expediting action; objective or subjective evaluations may not have the power to do so. "We must in all cases distinguish between what I say or mean, why I said it, and what effect it will have on my listener" (Baier, 1958, page 4). All evaluators share this need. Ayer (1959, page 228) asks:

> What is it that determines the truth or falsehood of empirical proposition? The customary answer is, in effect, that it is their agreement or disagreement with reality... This answer, though I believe it to be correct, requires some elucidation. To quote William James: "Pragmatists and Intellectualists both accept (it) as a matter of course. They begin to quarrel only after the question is raised as to what precisely may be meant by the term 'agreement' and what by the term 'reality' when reality is taken as something for our ideas to agree with."

The emotive-imperative evaluator stresses that disagreements in educational policy occur when attitudes and feelings cannot be satisfied. Reasoning empirically will often reduce conflict; often it does not. Educational decisions are very difficult when made under conditions of controversy. At this point, it may be advantageous for the policy maker or program director to initiate an emotive-imperative evaluation attempt. Specific emotions are frequently associated with actual educational activities, although this existence may be very fragile. While the objectivist can identify the primary qualities of educational programs, the possible emotional attachments to programs may return to haunt the educational policy maker.

The rationale of the emotive-imperative evaluator is the presentation of results based on beliefs about the educational program examined and the desire to arouse similar beliefs and feelings from policy makers. Frequently, such evaluations are designed to garner support for an educational service or to further educational philosophies of individuals or groups.

Instrumentalism

To discuss instrumental evaluation we must first return to our two educational questions: (1) "What things ought to exist in edcaution?" (the ends of education) and (2) "What actions best bring about what should exist?" (the means of education). Here "best educational actions" are evaluated as educational means which further some educational end; things are judged as good when they lead to other good things. Instrumentalism is a method of making decisions which resolve problems in practical situations. Therefore, an instrumental evaluation cannot be understood or interpreted outside the context of the specific educational program being evaluated. Its value to the policy maker is exact and predefined.

The theory of instrumentalism was first presented by John Dewey as a meta-ethical theory. Basically, this process involves the use of scientific validation to generate answers for specific questions. Dewey (1966) has described two types of educational values; intrinsic and instrumental.

> To value means primarily to prize, to esteem; but secondarily it means to apprize, to estimate. It means, that is, the act of cherishing something, holding it dear, and also the act of passing judgment upon the nature and amount of its value as compared with something else. To value in the latter sense is to valuate or evaluate. The distinction coincides with that sometimes made between intrinsin and instrumental values. Intrinsic values are not objects of judgment, they cannot (as instrumental) be compared, or regarded as greater and less, better or worse. They are invaluable; and if a thing is invaluable, it is neither more nor less so than any other invaluable. But occasions present themselves when it is necessary to choose, when we must let one thing go in order to take another. This establishes an order of preference, a greater and less, better and worse. Things judged or passed upon have to be estimated in relation to some third thing, some further end. With respect to that, they are means, or instrumental values (page 238).

Intrinsic values are held by every individual and no policy purpose is served by attempts to compare or evaluate these values. Instrumental values, on the other hand, compete for attention and become contestants as the means toward further educational ends. Most would argue that an education is invaluable; its worth in our society is generally known. However, the implementation of educational programs involves many fundamental policy decisions; the evaluator seeks to find what is most instrumnetal in accomplishing specified educational ends. A distinct advantage of instrumental evaluation is the preciseness and unambiguity of evaluative rationale. Terms such as "successful" or "failure" have no meaning; there is no need for debate over such general conclusions. Instrumental evaluation studies specific variables, and research findings are directly attributable to those variables. An educational program was instrumental in raising achievement scores; it was not instrumental, perhaps, in changing attitudes toward schooling. This procedure helps alleviate situations where educational programs are simultaneously evaluated by various sources to be "good" and "bad," and helps reduce the confusion caused the policy maker in such situations.

The term instrumentality replaces many (actually all other) labels now commonly assigned as descriptions of educational programs. A parallel is drawn here between instrumentality to evaluation and generalizability to reliability as discussed by Cronbach (1970, page 154):

> When we use a single observation as if it represented the universe, we are generalizing... If the observed scores from a procedure agree closely with the universe score, we can say that the observation is "accurate," or "reliable," or "generalizable." ...To have so many alternative terms is confusing, but not seriously so. The term most often used in the literature is "reliability." The author prefers "generalizability" because that term immediately implies "generalization to what?"

The determining of the value of a policy is a less clear process than finding the instrumentality of a policy to do that which it was designed to do; in instrumental evaluation the question immediately follows: "Instrumental in what way."

There is nothing prescriptive about instrumental evaluation; the problematic situation must be available or predicted before such evaluation. Nothing is final when dealing with instrumental values--there is a circular dependency on the educational means leading to ends which lead to other means in a never ending cycle. Evaluations of this type are associated with the operations performed during the implementation of educational programs or policy. "Prior experience supplies the conditions which evoke ideas and of which thought has to take account, with which it must reckon. It furnishes both obstacles to attainment of what is desired and the resources that must be used to attain it" (Gouinlock, 1976, page 122).

Dewey has not been without critics. A common criticism of instrumentalism is its preoccupation with educational practice and neglect of educational policy. The implementation of instrumental evaluations may not give direction to policy; thus, without policy for action, educational progress may be inhibited. The nebulous relationship between educational ends and means has also been questioned. "Or again, in general terms, it is one thing to assert that an end has a value because it is used or usable as a means, and quite another to identify ends with means. That Dewey imperceptibly passes from one position to the other is seen in many of his writings" (Sathaye, 1972, page 802).

The instrumental evaluator holds education as an on-going activity with constant problems requiring specific decisions. The instrumental evaluator works in specific educational contexts; generalizability or external validity of research findings may not be of direct interest to the policy maker. The application of scientific technique and thinking is paramount; educational policy is secondary to educational practice.

CONCLUSION

Of the four evaluation systems discussed, subjectivism and objectivism constitute the large majority of applications in educational evaluation. Public schools and programs are influenced by political values, public values, as well as educational values. Some values are so established they are taken for granted: the value of education itself, the value of public schooling, the value of teacher training. "Among the diverse values held by Americans, the idea of progress is uniquely important; first because it can be viewed not only as an individual, but also as an institutional, and even a national value; and second, because other values are subordinate to it" (Burke, 1969, page 200).

When discussions of evaluation systems do occur they are frequently on a shallow level; the differences between objectivism and subjectivism may be reduced to arguments over quantitative data versus qualitative data. Qualitative data are often treated apologetically (since it is "unreliable, biased, opinionated, etc.") while quantitative has come to mean reliable, confirmable and factual. Actually, it is as possible to test subjective areas as objective areas. The difference lies in the difficulty of collecting data; however, evaluation research is not concerned only with what is easily confirmable. The use of objective data is a subjective decision in itself--no evaluator fits in only one of the four evaluation systems. Any one of these systems may become an ideology which dictates methodological procedure. For instance, the objectivist may ignore the reality or importance of some educational variables if objective data are not available concerning those variables.

The rationale of any evaluation subsumes at least one of these four evaluation systems. Identification of the system(s) can aid in the interpretation of reported results, can be used as a measure of the usefulness of the study, and can sort evaluation studies into larger collections of data which serve as a base for the establishment of educational policy. "For the social sciences to refuse to treat their own behavior as data from which one can learn is really tragic" (Scriven, 1972, page 99).

REFERENCES

Ayer, A. J. Language, truth and logic. New York: Dover Publications, 1952.

Ayer, A. J. Verification and experience. In A. J. Ayer (Ed.), Logical positivism. Glencoe, Illinois: The Free Press, 1959.

Baier, K. The moral point of view: a rationale basis of ethics. New York: Cornell University Press, 1958.

Bissell, J. S. Planned variation in Head Start and Follow Through. In. J. C. Stanley (Ed.), Compensatory education for children ages two to eight: recent studies of educational intervention. Baltimore: The Johns Hopkins Press, 1973.

Blanshard, B. Reason and analysis. La Salle, Illinois: Open Court Publishing Company, 1962.

Blanshard, B. Reason and goodness. London: George Allen & Unwin Ltd., 1961.

Burke, J. G. Technology and values. In C. Bird (Ed.), The great ideas today 1969. New York: Praeger Publishing, Inc., 1969.

Byrne, E. F. Probability and opinion. The Hague: Martinus Nijhoff, 1968.

Campbell, D. T., & Stanley, J. C. Experimental and quasi-experimental designs for research. Chicago: Rand McNally College Publishing Company, 1963.

Caws, P. Science and the theory of value. New York: Random House, 1967.

Cronbach, L. J. Essentials of psychological testing. (3rd ed.) New York: Harper & Row, 1949.

Dewey, J. Democracy and education. New York: The Free Press, 1966.

Dewey, J. Theory of valuation. Chicago: University of Chicago Press, 1939.

Fitz-Gibbon, C. T., & Morris, L. L. Theory-based evaluation. Evaluation Comment, 1975, 5(1).

Frondizi, R. What is value? La Salle, Illinois: Open Court Publishing Company, 1971.

Gouinlock, J. The moral writings of John Dewey. New York: Hafner Press, 1976.

Guttentag, M. Subjectivity and its use in evaluation research. Evaluation, 1973, 3(1). 7-14

Jensen, A. R. How much can we boost IQ and scholastic achievement? Harvard Educational Review, 1969, 39(1), 1-123.

Kyburg, H. E., & Smokler, H. E. (Eds.). Studies in subjective probability. New York: John Wiley & Sons, Inc., 1964.

Lepley, R. Verifiability of value. New York: Columbia University Press, 1944.

Moore, G. E. Principia ethica. Cambridge: At the University Press, 1948.

Perry, R. B. General theory of value. New York: Longmans, Green and Company, 1926.

Perry, R. B. Realms of value: a critique of human civilization. New York: Greenwood Press, 1968.

Ross, W. D. The right and the good. Oxford: At the Clarendon Press, 1930.

Russell, B. The analysis of mind. London: George Allen & Unwin Ltd., 1924.

Sathaye, S. G. Instrumentalism: a methodological exposition of the philosophy of John Dewey. Bombay: Popular Prakshan, 1972.

Scriven, M. Objectivity and subjectivity in educational research. In L. G. Thomas (Ed.), Philosophical redirection of educational research. Chicago: The University of Chicago Press, 1972.

Stevenson, C. L. Ethics and language. New Haven: Yale University Press, 1944.

Stevenson, C. L. Facts and values: studies in ethical analysis. New Haven: Yale University Press, 1963.

Stufflebeam, D. L., et.al., Educational evaluation & decision making. Itasca, Illinois: F. E. Peacock Publishers, Inc., 1971.

Taylor, P. W. The moral judgment. Englewood Cliffs, New Jersey: Prentice-Hall, Inc., 1963.

Thorndike, E. L. Seventeenth yearbook of the National Society for the Study of Education, Part II, 1918.

Toulmin, S. E. An examination of the place of reason in ethics. Cambridge: Cambridge University Press, 1950.

Westermarck, E. A. Ethical relativity. Paterson, New Jersey: Littlefield, Adams, 1960.

White, M. G. The origins of Dewey's instrumentalism. New York: Octagon Books, Inc., 1964.